

Extending the Degree Method in Complexity Theory

Kenneth W. Regan

`regan@cse.buffalo.edu`

Department of Computer Science and Engineering
State University of New York at Buffalo
Buffalo, NY 14260 USA

The difficulty with lower bounds

Arithmetical or Boolean circuit size $s(f)$ does not correspond simply to a known mathematical quantity $q(f)$ with a beautiful pedigree.

Here $f = f_n : F^n \longrightarrow F$ where the field F can be $\{0, 1\}$ or \mathbb{Q} or \mathbb{R} or \mathbb{C} , etc.

Strategy

Find a good $q(f)$ and a simple function ℓ and prove

$$s(f) \geq \ell(q(f)).$$

Strategy

Find a good $q(f)$ and a simple function ℓ and prove

$$s(f) \geq \ell(q(f)).$$

Strassen (+Baur 1983; OSW 1971):

$$s(f) \geq \frac{1}{3} \log_2(\text{gdeg}(\text{graph}(\partial f))).$$

Strategy

Find a good $q(f)$ and a simple function ℓ and prove

$$s(f) \geq \ell(q(f)).$$

Strassen (+Baur 1983; OSW 1971):

$$s(f) \geq \frac{1}{3} \log_2(\text{gdeg}(\text{graph}(\partial f))).$$

Here f is a multi-variate polynomial, and

$s(f)$: arithmetical complexity over F , an *infinite* field

∂f : the list $\partial f / \partial x_1, \dots, \partial f / \partial x_n$

graph: creates $y_1 - \partial f / \partial x_1, \dots, y_n - \partial f / \partial x_n$

gdeg: *geometric degree* of the graph.

Arithmetical Circuits C

- DAG with gates $+$ and $*$ (no $/$), binary unless stated otherwise
- Multiplicative scalars in F allowed on wires, so $a - b = a + (-1)b$.
- Input nodes x_1, \dots, x_n .
- Output nodes g_1, \dots, g_m —need not be sinks.

If $m > 1$, C computes a *regular function*

$$\vec{f} : F^n \longrightarrow F^m, \vec{f} = (f_1, \dots, f_m).$$

We try to avoid source nodes with constants—unless we really need to compute e.g. $f(x) = x + 3$.

C is (hereditarily) homogeneous if every $+$ gate has inputs of the same *formal degree*.

C Is-A formula if every gate has fan-out 1.

Choose your field!

Challenge: Compute $f(x, y) = x^2 + y^2$ with:

Choose your field!

Challenge: Compute $f(x, y) = x^2 + y^2$ with:
one $*$ gate:

Choose your field!

Challenge: Compute $f(x, y) = x^2 + y^2$ with:
one * gate: $(x + iy) * (x - iy)$ ($F = \mathbb{C}$)

Choose your field!

Challenge: Compute $f(x, y) = x^2 + y^2$ with:

one $*$ gate: $(x + iy) * (x - iy)$ ($F = \mathbb{C}$)

no $*$ gates:

Choose your field!

Challenge: Compute $f(x, y) = x^2 + y^2$ with:

one * gate: $(x + iy) * (x - iy)$ ($F = \mathbb{C}$)

no * gates: $x + y$ ($F = \mathbb{Z}_2$)

(Can we find a field to do $x^2 + y^2 + z^2 + w^2$ with one * gate?)

Choose your field!

Challenge: Compute $f(x, y) = x^2 + y^2$ with:

one * gate: $(x + iy) * (x - iy)$ ($F = \mathbb{C}$)

no * gates: $x + y$ ($F = \mathbb{Z}_2$)

(Can we find a field to do $x^2 + y^2 + z^2 + w^2$ with one * gate?)

Ontology: A polynomial function f is named by some formula ϕ or circuit C with scalars in some “base field” F_0 , and computes “the same” f over all *extension fields* $F : F_0$.

Half of the theory doesn't care what the field F is!
Much of the other half cares only about:

- is F finite or infinite?
- if finite, is its characteristic 2 or an odd prime?
- if infinite, is its characteristic 0?
- is F *algebraically closed*?

Circuit Complexity Measures

$s_+(C)$: number of $+$ gates in C

$s_*(C)$: number of $*$ gates

$s(C) := s_+(C) + s_*(C)$ (input nodes not counted)

$w(C)$: number of wires in C , $= 2s(C)$ for binary circuits

$\ell(\phi) =$ number of *leaves* in formula ϕ , $= 1 + s(\phi)$

For f defined over an F_0 ,

$$s(f) = \min\{ s(C) : (\exists F : F_0) C = f \text{ over } F \}$$

$$\ell(f) = \min\{ \ell(\phi) : (\exists F : F_0) \text{formula } \phi = f \text{ over } F \}.$$

Example: for $f = x^2 + y^2$, named over \mathbb{Q} ,

$$s_*(f) = \ell_*(f) = 1.$$

Point: lower bounds on “ $s(f)$ ” apply to all $F : F_0$. We can still write $s_*^{\mathbb{R}}(f) = 2$.

For f defined over an F_0 ,

$$s(f) = \min\{ s(C) : (\exists F : F_0) C = f \text{ over } F \}$$

$$\ell(f) = \min\{ \ell(\phi) : (\exists F : F_0) \text{formula } \phi = f \text{ over } F \}.$$

Example: for $f = x^2 + y^2$, named over \mathbb{Q} ,

$$s_*(f) = \ell_*(f) = 1.$$

Point: lower bounds on “ $s(f)$ ” apply to all $F : F_0$. We can still write $s_*^{\mathbb{R}}(f) = 2$. Extra super/subscripts can restrict to families of circuits, such as

$$s^h(f) = \text{min size of a } \textit{homogeneous} \text{ circuit computing } f.$$

Geometry of polynomials

For a polynomial $f \in F[x_1, \dots, x_n]$, and list B of such polynomials,

$$V_f = \{ a \in F^n : f(a) = 0 \} \text{ (hypersurface if irreducible)}$$

$$V_{B,f} = V_B \cap V_f \text{ (algebraic set, aka. variety? yes if irreducible)}$$

$$V_0 = F^n, V_1 = \emptyset.$$

V_L , for $C = x^2 + y^2 - 1$ and $L = x - y = 1/2$ in \mathbb{R}^2 , and their intersection.

V is *irreducible* if V cannot be written as the union of two *other* algebraic sets. Note $V_{f*g} = V_f \cup V_g$.

Zariski Topology

The *Zariski topology on F^n* has (irreducible) algebraic sets as its (basic) closed sets.

Zariski Topology

The *Zariski topology* on F^n has (irreducible) algebraic sets as its (basic) closed sets.

If F is finite, every subset of F^n is algebraic, so Z is trivial.

For $F : \mathbb{R}^n$, every Z -closed set is Euclidean-closed, but not conversely: all except V_0 have zero Lebesgue measure, so every Z -open set except V_1 is dense.

Zariski Topology

The *Zariski topology on F^n* has (irreducible) algebraic sets as its (basic) closed sets.

If F is finite, every subset of F^n is algebraic, so Z is trivial.

For $F : \mathbb{R}^n$, every Z -closed set is Euclidean-closed, but not conversely: all except V_0 have zero Lebesgue measure, so every Z -open set except V_1 is dense.

Restricting to homogeneous polynomials $f(x_0, x_1, \dots, x_n)$ defines the Z -topology on *projective space* \mathbb{P}^n .

Geometric Degree—affine

Irreducible V have a well-defined *dimension* d .

Define $gdeg(V)$ to be the maximum size of a finite intersection of V with an affine-linear subspace of dimension $n - d$.

Geometric Degree—affine

Irreducible V have a well-defined *dimension* d .
Define $gdeg(V)$ to be the maximum size of a finite intersection of V with an affine-linear subspace of dimension $n - d$.

Example:

$$\dim(V_C) = 1, gdeg(V_C) = 2.$$

The algebraic set $S = V_C \cup V_L = V_{C*L}$ is reducible but still *equidimensional*, so you can say $\dim(S) = 1$ (glatt-kosher) and $gdeg(S) = 3$ (merely kosher).

Some properties

- The graph of a regular function $\vec{f} : F^n \longrightarrow F^m$ is always irreducible in F^{n+m} .
- $gdeg(V_{y-f}) = \deg(f)$, so $gdeg$ extends the ordinary notion of the degree of (the graph of) a polynomial.
- Geometric degree does not increase under any kind of projection.
- So $gdeg(V_f) \leq gdeg(V_{y-f})$ (still kosher?)
- Attention! $gdeg(V_{f^2}) = gdeg(V_f)$. So $gdeg(V_{x^n}) = 1$, all n .
- For a finite set S of points, $\dim(S) = 0$ and $gdeg(S) = |S|$.

Bezout Inequality, affine

Lemma 0.1 *If $gdeg(V_B)$ is defined and V_f is a hypersurface, then $gdeg(V_{B,f}) \leq gdeg(V_B) \cdot \deg(f)$.*

Bezout Inequality, affine

Lemma 0.2 *If $gdeg(V_B)$ is defined and V_f is a hypersurface, then $gdeg(V_{B,f}) \leq gdeg(V_B) \cdot \deg(f)$.*

Question: Can we prove this without geometry, with the “unkosher” definition and statement:

$udeg(f_1, \dots, f_k) = \max\{r \in \mathbb{N} : \text{one can adjoin } n - k \text{ affine linear equations to make a system with exactly } r \text{ solutions}\}.$

$$udeg(f_1, \dots, f_k) \leq udeg(f_1, \dots, f_{k-1}) \cdot \deg(f_k),$$

not caring about irreducibles, and with high-school algebra? (Bezout’s **Theorem** counts multiplicities and gives $gdeg(\cdot) = \dots$; he proved it in the Napoleonic era using resultants.)

Strassen's Degree Bound

Theorem 0.3 (Str72) *For any regular function \vec{f} ,
 $s_*(\vec{f}) \geq \log_2(\text{gdeg}(\text{graph}(\vec{f})))$.*

Strassen's Degree Bound

Theorem 0.4 (Str72) For any regular function \vec{f} ,
 $s_*(\vec{f}) \geq \log_2(\text{gdeg}(\text{graph}(\vec{f})))$.

Proof. Let C compute \vec{f} over (your favorite) F . For each node i let g_i name the polynomial it computes and give a new variable y_i . Then $\vec{G} := \{y_i - g_i\}$ is a set of irreducibles that includes $\text{graph}(\vec{f})$, so $\text{gdeg}(\vec{G}) \geq q = \text{gdeg}(\text{graph}(\vec{f}))$.

Rest of proof

Now an equivalent system to \vec{G} is defined as follows: Start with $\{y_1 - x_1, \dots, y_n - x_n\}$. For each $+$ gate $g_k = cg_i + dg_j$ (here $k > n$ and $k > j \geq i \geq 1$), adjoin $y_k - cy_i - dy_j$. For each $*$ gate $g_k = cg_i g_j$, adjoin $y_k - cy_i y_j$. Then each polynomial h_k thus adjoined is irreducible, so the Bezout inequality gives

$$gdeg(\{h_k\}) \leq \prod_k (\deg(h_k)) = 2^{s_*(C)}.$$

Finally $V_{\vec{G}} = V_{\{h_k\}}$. So

$$s_*(C) \geq \log_2(gdeg(\{h_k\})) = \log_2(gdeg(\vec{G})) \geq q. \quad \square$$

For a single polynomial

Objection: \vec{f} is not a single polynomial.

Lemma 0.5 (“Derivative Lemma” OSW71, BaSt82)

$$s(f) = \Theta(s(f, \partial f)).$$

Proof

Let C compute f . Topologically number its nodes.

For each node g_j , define $f^{(j)}(x_1, \dots, x_n, y_j) =$ the function computed when node/gate g_j is replaced by the new input node y_j . Formally define

$$\frac{\partial f}{\partial g_j} = \frac{\partial f^{(j)}}{\partial y_j}.$$

For input nodes x_j this $= \partial f / \partial x_j$, and for the output gate g_{n+s} , $\partial f / \partial g_{n+s} = 1$.

Now the nodes h_1, \dots, h_k that g_j fans out to each have y_j as an input in the circuit for $f^{(j)}$. Thus by the Chain Rule,

$$\frac{\partial f^{(j)}}{\partial y_j} = \sum_{i=1}^k \frac{\partial f^{(j)}}{\partial h_k} * \frac{\partial h_k}{\partial y_j}$$

So

$$\frac{\partial f}{\partial g_j} = \sum_{i=1}^k \frac{\partial f}{\partial h_k} * \frac{\partial h_k}{\partial g_j}$$

Proof, continued

$$\frac{\partial f}{\partial g_j} = \sum_{i=1}^k \frac{\partial f}{\partial h_k} * \frac{\partial h_k}{\partial g_j}$$

By topo order and induction we've already computed the first terms in the sum. For a $+$ gate h_k , the latter term is a constant, so no $*$ is needed. For a $*$ gate h_k , $\partial h_k / \partial g_j$ is a constant times the other input to h_k , which we already have. Thus we add:

$k - 1$ addition gates

At most k multiplication gates.

So in particular,

$$s_*(f, \partial f) \leq s_*(f) + w_*(C) \leq 3s_*(C). \quad \square \square$$

Attention: This uses a gate constant 1.

If C ends with m gates that iteratively sum each other, it builds a gate constant 2^m .

Thus the Derivative Lemma does not preserve the bc-property.

Nor does it preserve being a formula. It *does* preserve homogeneity.

Example

$$f = x_1^d + \dots + x_n^d$$

Example

$$f = x_1^d + \dots + x_n^d$$

$$\text{graph}(\partial f) = (y_1 - dx_1^{d-1}, \dots, y_n - dx_n^{d-1})$$

Example

$$f = x_1^d + \dots + x_n^d$$

$$\text{graph}(\partial f) = (y_1 - dx_1^{d-1}, \dots, y_n - dx_n^{d-1})$$

Adjoin affine linear $(y_1 = 1, \dots, y_n = 1)$

Example

$$f = x_1^d + \dots + x_n^d$$

$$\text{graph}(\partial f) = (y_1 - dx_1^{d-1}, \dots, y_n - dx_n^{d-1})$$

Adjoin affine linear $(y_1 = 1, \dots, y_n = 1)$

Over \mathbb{C} the system has $(d - 1)^n$ solutions

Example

$$f = x_1^d + \dots + x_n^d$$

$$\text{graph}(\partial f) = (y_1 - dx_1^{d-1}, \dots, y_n - dx_n^{d-1})$$

Adjoin affine linear $(y_1 = 1, \dots, y_n = 1)$

Over \mathbb{C} the system has $(d - 1)^n$ solutions

$$\text{So } s_*(f) = \Omega(\log((d - 1)^n)) = \Omega(n \log d)$$

Example

$$f = x_1^d + \dots + x_n^d$$

$$\text{graph}(\partial f) = (y_1 - dx_1^{d-1}, \dots, y_n - dx_n^{d-1})$$

Adjoin affine linear $(y_1 = 1, \dots, y_n = 1)$

Over \mathbb{C} the system has $(d - 1)^n$ solutions

$$\text{So } s_*(f) = \Omega(\log((d - 1)^n)) = \Omega(n \log d)$$

Which $= \Omega(n \log n)$ when $d = n$, and is tight for f .

The limitation

“Bezout’s Double-Edged Sword”: Applied to gate equations,

$$gdeg(C) \leq 2^{s_*(C)}.$$

But B.I. applies directly to (f_1, \dots, f_N) to get

$$gdeg \leq \prod_i \deg(f_i).$$

If $\deg(f) = d + 1$, then $\partial^1 f$ gives $N = n$ and all $\deg(f_i) = d$, so

$$gdeg \leq d^n.$$

Thus for $d = n^{O(1)}$, $s(C) = \Omega(n \log n)$ is the best you can do.

Overcoming the limitation

1. Transform to cases where $d \gg n^{O(1)}$

Overcoming the limitation

1. Transform to cases where $d \gg n^{O(1)}$
2. Implicitly handle N -tuples (f_1, \dots, f_N) where $N \gg n$

Overcoming the limitation

1. Transform to cases where $d \gg n^{O(1)}$
2. Implicitly handle N -tuples (f_1, \dots, f_N) where $N \gg n$
3. For restricted circuit classes, get tighter gate equations

Overcoming the limitation

1. Transform to cases where $d \gg n^{O(1)}$
2. Implicitly handle N -tuples (f_1, \dots, f_N) where $N \gg n$
3. For restricted circuit classes, get tighter gate equations
4. Find a q that gives $s \geq \log(q)$, can reach magnitude (say) d^{2^n} , and is *not* sub-multiplicative on tuples!

Overcoming the limitation

1. Transform to cases where $d \gg n^{O(1)}$
2. Implicitly handle N -tuples (f_1, \dots, f_N) where $N \gg n$
3. For restricted circuit classes, get tighter gate equations
4. Find a q that gives $s \geq \log(q)$, can reach magnitude (say) d^{2^n} , and is *not* sub-multiplicative on tuples!
5. Do something else.

Concrete ideas...

1. Good relations between VP,VNP and WP,WNP?

Concrete ideas...

1. Good relations between VP,VNP and WP,WNP?
2. Higher partial derivatives, with an iterable Derivative Lemma...

Concrete ideas...

1. Good relations between VP,VNP and WP,WNP?
2. Higher partial derivatives, with an iterable Derivative Lemma...
3. Hereditarily multilinear formulas...

Concrete ideas...

1. Good relations between VP,VNP and WP,WNP?
2. Higher partial derivatives, with an iterable Derivative Lemma...
3. Hereditarily multilinear formulas...
4. A candidate, alas refuted: $q(f_1, \dots, f_N) =$ the number of monomials m such that $m \in I = \{ \sum_k p_k f_k : p_1, \dots, p_N \in F[x_1, \dots, x_n] \}$ (the *ideal* generated by f_1, \dots, f_N), but no proper divisor of m is in I .

Concrete ideas...

1. Good relations between VP,VNP and WP,WNP?
2. Higher partial derivatives, with an iterable Derivative Lemma...
3. Hereditarily multilinear formulas...
4. A candidate, alas refuted: $q(f_1, \dots, f_N) =$ the number of monomials m such that $m \in I = \{ \sum_k p_k f_k : p_1, \dots, p_N \in F[x_1, \dots, x_n] \}$ (the *ideal* generated by f_1, \dots, f_N), but no proper divisor of m is in I .
5. Mulmuley-Sohoni's use of representation theory and (partial) *stability*.